REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
DoDPI01-R-0008

3. REPORT TYPE AND DATES COVERED
Final Report (June 1997 - May 2001)

4. TITLE AND SUBTITLE
Effects of Deception on Tonic Autonomic Arousal

5. FUNDING NUMBERS
DoDPI97-P-0016

6. AUTHOR(S)
John C. Kricher, Ph.D.; Ted Packard, Ph.D.; Brian G. Bell, Ph.D.; Pauld C. Bernahardt, M.S.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Educational Psychology
University of Utah
Salt Lake City, UT 84112

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DoD Polygraph Institute
7540 Pickens Avenue
Fort Jackson, SC 29207

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
DoDPI97-P-0016
DoDPI01-R-0008

11. SUPPLEMENTARY NOTES

20011203 190

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Public release, distribution unlimited

A

13. ABSTRACT *(Maximum 200 words)*
The present study tested if measures of tonic arousal are related to the amplitude of responses during probable-lie and directed lie polygraph examinations. It also tested if tonic levels of electrodermal and cardiovascular activity can be used to improve the accuracy of polygraph examinations. Three hundred and thirty-six male and female participants in a previous experiment (DODPI97-P-0016) were interrogated about their participation in a mock crime. Half of the subjects were guilty of committing the mock crime and half were innocent. Half of the innocent subjects and half of the guilty subjects received a probable-lie polygraph examination. The remaining subjects received a directed lie polygraph examination. Subjects were offered a $50 bonus to appear truthful on the polygraph examination. Tonic and phasic measures of skin conductance and skin resistance were obtained from skin conductance recordings. Tonic and phasic systolic and diastolic blood pressure measures were obtained from a Finapresâ blood pressure monitor. Tonic and phasic measures of heart period were obtained from the electrocardiogram (EKG).
Consistent with prior research, two tonic measures of skin conductance were positively related to phasic reactivity. Weak but significant correlations between tonic and phasic arousal also were obtained for blood pressure and heart period. However, none of the tonic measures improved the accuracy of polygraph outcomes. Tonic arousal accounted for less than 2% of the variance in the guilt/innocence criterion when used in combination with standard measures of differential reactivity to predict group membership. The results suggest that the use of absolute measures of electrodermal and cardiovascular activity would do little to improve the accuracy of computer algorithms for diagnosing truth and deception.

14. SUBJECT TERMS

15. NUMBER OF PAGES
30

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT

DoDPI01-R-0008

Effects of Deception on Tonic Autonomic Arousal

John C. Kircher, Ph.D., Ted Packard, Ph.D., Brian G. Bell, Ph.D., Paul C. Bernhardt, M.S.
University of Utah, Salt Lake City, UT 84112

11 May, 2001

Department of Defense Polygraph Institute
7540 Pickens Avenue
Fort Jackson, SC 29207

Final Report

Project Title: *Effects of Deception on Tonic Autonomic Arousal*

Date: 14 March, 2000
Revised: 11 May, 2001

Submitted by: John C. Kircher, Ph.D. (Principal Investigator)
Department of Educational Psychology
University of Utah
Salt Lake City, UT 84112
voice: (801) 581-7130
fax: (801) 581-5566
email: kircher@gse.utah.edu

Ted Packard, Ph.D. (Co-principal Investigator)
Department of Educational Psychology
University of Utah
Salt Lake City, UT 84112
voice: (801) 278-5958
fax: (801) 581-5566
email: packard@gse.utah.edu

Brian G. Bell, Ph.D. (Research Associate)
Department of Educational Psychology
University of Utah
Salt Lake City, UT 84112

Paul C. Bernhardt, M.S. (Graduate Assistant)
Department of Educational Psychology
University of Utah
Salt Lake City, UT 84112

Director's Foreword

Psychophysiological detection of deception research is constantly striving to advance and perfect procedures and techniques to optimize detection accuracy. In doing so, a number of processes and data sources must be considered. One strong area in this regard is the examination of new and untapped physiological data sources that may be diagnostic in the detection of deception. This project examined the diagnostic contributions of a data source that has typically been ignored and eliminated by most computer algorithms used to produce decisions on polygraph data, that of tonic electrodermal and cardiovascular arousal. At least two manufacturers of computer algorithms detrend and effectively eliminate these tonic responses, instead focusing upon the analysis of phasic responses. The results of this study indicate that tonic sources of information from electrodermal and cardiovascular channels are of little or no diagnostic value for the detection of deception. To the extent that the results of this laboratory study can be compared to field situations, there is no evidence from this study suggesting that manufacturers of computer algorithms should incorporate tonic electrodermal and/or cardiovascular arousal into their evaluative processes. Future research should continue to determine the diagnostic value of new and unexplored data sources.

William F. Norris

William F. Norris
Director

Abstract

The present study tested if measures of tonic arousal are related to the amplitude of responses during probable-lie and directed lie polygraph examinations. It also tested if tonic levels of electrodermal and cardiovascular activity can be used to improve the accuracy of polygraph examinations. Three hundred and thirty-six male and female participants in a previous experiment (DODPI97-P-0016) were interrogated about their participation in a mock crime. Half of the subjects were guilty of committing the mock crime and half were innocent. Half of the innocent subjects and half of the guilty subjects received a probable-lie polygraph examination. The remaining subjects received a directed lie polygraph examination. Subjects were offered a $50 bonus to appear truthful on the polygraph examination. Tonic and phasic measures of skin conductance and skin resistance were obtained from skin conductance recordings. Tonic and phasic systolic and diastolic blood pressure measures were obtained from a Finapres® blood pressure monitor. Tonic and phasic measures of heart period were obtained from the electrocardiogram (EKG).

Consistent with prior research, two tonic measures of skin conductance were positively related to phasic reactivity. Weak but significant correlations between tonic and phasic arousal also were obtained for blood pressure and heart period. However, none of the tonic measures improved the accuracy of polygraph outcomes. Tonic arousal accounted for less than 2% of the variance in the guilt/innocence criterion when used in combination with standard measures of differential reactivity to predict group membership. The results suggest that the use of absolute measures of electrodermal and cardiovascular activity would do little to improve the accuracy of computer algorithms for diagnosing truth and deception.

Introduction

The primary objective of the present study was to determine if tonic levels of skin conductance and other measures of tonic arousal can be used to improve the accuracy of probable-lie and directed lie polygraph tests. The present study also assessed the extent to which tonic skin conductance and other measures of tonic arousal (skin resistance, arterial blood pressure, and heart period) are related to the magnitude of responses during probable-lie and directed lie polygraph examinations.

Background

It is well known that there are large differences among individuals in both tonic levels of physiological arousal and in the magnitude of phasic responses to stimuli. Some polygraph subjects will have basal skin conductance levels that measure less than one μSiemen (1 M ohm), whereas other subjects in the same situation will have skin conductance levels that approach 100 μSiemens (10 K ohm; Venables & Christie, 1980). Nonspecific skin conductance responses, another measure of tonic arousal, may range from zero to 10 per minute (Boucsein, 1992). Tonic heart rate ranges from 50 BPM to over 120 BPM, and tonic levels of mean arterial blood pressure range from 70 mm Hg to over 130 mm Hg (Rushmer, 1976).

Computer and numerical scoring procedures remove individual differences in tonic arousal by making within-subject comparisons of the individual's physiological reactions to different types of test questions. If a reaction is noticeably greater to one question than to another, the larger reaction is considered diagnostic (Raskin, 1989). To make these judgments, the computer or polygraph examiner considers only the relative strength of reactions to different types of test questions. No systematic attempt is made to account for the fact that reactions to test questions are superimposed on a baseline of tonic activity.

Psychophysiological research indicates that tonic levels of activation correlate with the magnitude of evoked responses to stimuli (e.g., Hord, Johnson, & Lubin, 1964). Tonic arousal also predicts the habituation of responses that typically occurs with repeated presentations of a stimulus. Katkin (1975) found that subjects with high levels of electrodermal activity showed less habituation to a series of tones than did subjects with low levels of tonic electrodermal activity.

However, results from studies are mixed in which tonic levels were related to differential responses to signal and nonsignal stimuli. Some studies suggest that tonic levels predict differential responses to signal and nonsignal stimuli (Katkin, 1975), whereas others do not (Schell, Dawson, & Filion, 1988). The relationship between tonic arousal and differential reactivity to signal and nonsignal stimuli is important because polygraph examiners base their decisions on differences between reactions to test questions that differ in signal value (Raskin, 1979). The relationship between tonic arousal and habituation is important because a polygraph examiner may present the same basic question as many as 15 times over the course of a polygraph examination, and the subject's physiological responses tend to habituate. If measures of tonic arousal predict individuals' patterns of response during a polygraph examination, they might be used as statistical 'filters' to remove noise from physiological measures that are used to predict the criterion. In so doing, tonic arousal may enhance the ability of those measures to discriminate between truthful and deceptive subjects.

The possibility that individual differences in tonic arousal can be used to improve the accuracy of polygraph tests has never been investigated. The issue remains unexplored primarily because, with the exception of heart rate, traditional polygraph instrumentation provides only relative measures of change in physiological activity. Absolute measures of physiological activity are needed to determine if individual differences in tonic arousal can be used to increase decision accuracy.

The CPS-LAB system used to collect the polygraph charts for the present study allowed us to measure skin conductance in microSiemens (μS), skin resistance in ohms, heart period in ms, and blood pressure in mm Hg. The availability of absolute measures of tonic activity allowed us to investigate the possibility that this large untapped resource of physiological information might be used to increase the accuracy of probable-lie and directed lie tests.

The goals of the present study were twofold. First, based on research by Hord et al. (1964), we predicted that tonic levels of autonomic arousal would be related to the strength of phasic reactions to test questions observed during probable-lie and directed lie examinations. The second objective was to determine if tonic autonomic arousal could be used to improve the accuracy of probable-lie and directed lie tests.

Methods

Subjects

Four-hundred-and-seventeen adults were recruited from the general community by newspaper advertisements for a study that examined the effects of the demonstration test on the accuracy of probable-lie and directed lie polygraph examinations (DoDPI97-P-0016). The advertisements offered $30 for two hours of participation and the opportunity to earn an additional $50 bonus. Of the 417 individuals, 81 were eliminated from the study for a variety of reasons. Thirty-three subjects assigned to the guilty condition (16%) declined to participate after they received their instructions to commit a simulated theft. Eighteen individuals failed to follow instructions (e.g., did not commit the theft yet reported for their polygraph, arrived late, brought a child with them to the lab). Thirteen individuals were dismissed due to health problems. Health problems included reports of pain, less than 4 hours of sleep, and high blood pressure. Nine individuals assigned to the guilty condition (5%) confessed. Equipment problems and experimenter errors resulted in the loss of eight other individuals. The remaining 168 innocent and 168 guilty subjects were retained to fill the cells of the design matrix (described below).

The mean age of the sample was 30.7 years (SD = 11). Years of education ranged from 9 to 25 (M = 14.3, SD = 2.5). Most participants were Caucasian (87.5%), 5.7% were Hispanic, and the remaining 6.8% were Black, Asian, American Indian, or chose not to respond. Fifty-three percent of the participants were single, 33.9% were married, and the remaining 12.2% were divorced, separated, or widowed. Although a wide range of occupations was represented, over

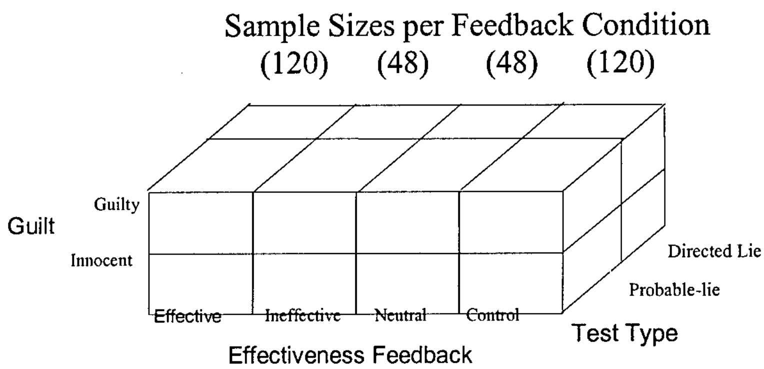
75% of the sample fell into one of the following eight categories: student (17%), professional (11.9%), sales worker (9.2%), office worker (8.3%), service worker (8.3%), unemployed (7.7%), homemaker (7.7%), or laborer (7.4%).

Design

Guilty and innocent subjects were randomly assigned to one of 16 cells in a completely crossed 2 x 2 x 4 factorial design with equal numbers of male and female participants in each cell. The design is illustrated in Figure 1. All factors except Sex are represented in the figure.

The first factor, Guilt, had two levels; 168 participants were guilty of committing a mock crime and the remaining 168 were innocent of the crime. The second factor, Test Type, also had two levels; half of the participants were given probable-lie comparison question tests (PL) and half were given directed lie tests (DL).

Figure 1. Design of experiment.



The third factor, Effectiveness Feedback, had four levels. Participants were unevenly distributed over the four levels of the Effectiveness Feedback factor. One group of 120 participants (30 participants in each of the four cells shown on the far left of Figure 1) received the type of feedback commonly provided to subjects in actual field examinations. Prior to their polygraph test, they were given a demonstration test and told, regardless of the outcome, that they showed their strongest reaction to the number they had chosen. They also were told they should have no problem passing the polygraph test if they were completely truthful to all of the questions (Effective Feedback group).

Twelve participants were assigned to each of the four Ineffective Feedback cells of the design matrix. Participants who received ineffective feedback were given a numbers test and were told, regardless of the outcome, that they did not react appropriately to the chosen number. They also were told that it would be difficult to determine if they were lying or telling the truth during their polygraph test.

Twelve participants were assigned to each of the four Neutral Feedback cells of the matrix. Participants who received neutral feedback were given a numbers test and were told that the test would provide an opportunity for the participant to practice answering questions and for the polygraph examiner to adjust the instrument. Participants were given no information about the outcome of the numbers test.

Thirty participants were assigned to each of the four control groups illustrated on the far right of Figure 1. The pretest procedures for subjects in the control groups were the same as those used for other subjects except that control subjects were not given a numbers test.

To summarize, 120 participants were given the demonstration test and received feedback that the test was effective. Another 48 participants were given a demonstration test and received feedback that the test was ineffective. Another 48 participants were given a demonstration test and received neutral feedback. The remaining 120 participants were not given a numbers test. Within each level of the Feedback factor, the design was balanced in terms of numbers of guilty and innocent male and female subjects who were given either probable-lie or directed lie polygraph examinations.

Two examiners administered all of the polygraph examinations. One examiner was an advanced doctoral student in educational psychology. The graduate student (PCB) tested 12 subjects in each of the 16 cells in the design matrix (192 subjects). The post-doctoral research associate (BGB) tested the remaining 144 subjects. The principal investigator (PI) trained both of the examiners. The PI has been conducting research on polygraph techniques at the University of Utah for the past 24 years and participated in annual workshops to train professional polygraph examiners for 17 of those years.

Procedures

The procedures followed those described elsewhere (Kircher & Raskin, 1988). Prospective participants called a secretary who screened the participants for eligibility and briefly described the experiment and pay policy. Callers were invited to participate if they met the following criteria: (1) they were between 18 and 65, (2) they were not taking prescription medication, (3) they had never had a polygraph test, (4) they were fluent in English, and (5) they had no major medical problems.

Callers who agreed to participate were given an appointment to report to a room in a building on the campus of the University of Utah. When the participant arrived, an envelope addressed to the participant was taped to the door. Instructions within the envelope directed the participant to enter the room, close the door, read and sign an informed consent form, complete a brief questionnaire, and then play a cassette recorder that presented their instructions over headphones.

Guilty participants received tape-recorded instructions to commit a mock theft of a $20 bill from a wallet that was in a purse in a desk in a secretary's office. Participants went to a secretary's office on a different floor of the building where they asked the secretary where Dr. Mitchell's office was located. The secretary was actually a confederate in the experiment. The secretary responded that there was no Dr. Mitchell in the department. The participant thanked the secretary and left the office. The participant then waited in the hallway until the secretary left the office unattended (1-3 minutes), entered the office, searched the desk for the purse, and took the $20 bill from the wallet that was in the purse. Participants were instructed to conceal the $20 on their person and go to a room where they waited for the polygraph examiner. Guilty participants were instructed to prepare an alibi in case they were caught in the office. Innocent participants listened to a general description of the crime, left the area for 15 minutes, and went to a room where they waited for the polygraph examiner.

All participants were told that they would be given a polygraph test by a polygraph expert who did not know whether they had committed the theft. They were told that the examiner would use a computer to assist in the analysis of their polygraph charts, and if they could convince the polygraph examiner of their innocence, they would receive $80. They were also told that if they failed to convince the examiner of their innocence, they would only receive $30.

After the participant had reported to the waiting room, the polygraph examiner went to the room, introduced himself, and instructed the participant to go to the restroom and wash their hands with soap and warm water. The participant was then escorted to the lab where the examiner obtained some biographical information and attached the sensors to the participant. Standard field polygraph procedures were used, and the session was videotaped and audiotaped. The polygraph examiner described the role of the autonomic nervous system in the detection of deception. He then described and administered the numbers test to participants in the effective, ineffective, and neutral feedback conditions. Finally, the polygraph examiner reviewed the appropriate set of test questions with the participant. Relevant questions that pertained to the theft and the sacrifice relevant question were reviewed first, probable-lie or directed lie comparison questions were reviewed next, and the neutral and outside issue questions were reviewed last. The test questions for participants assigned to the probable-lie condition were as follows:

(Outside Issue)	1. Do you understand that I will ask only the questions that we have discussed?
(Sacrifice Relevant)	2. Do you intend to answer truthfully all of the questions regarding the theft of the $20?
(Neutral)	3. Do you live in the United States?
(Probable-lie)	4. Before the age of ___, did you ever take something that didn't belong to you?
(Relevant)	5. Did you take that $20 from the purse?
(Neutral)	6. Is today ____?
(Probable-lie)	7. During the first ____ years of your life, did you ever do anything that was dishonest or illegal?
(Relevant)	8. Did you take that $20?
(Neutral)	9. Is your first name ___?

(Probable-lie) 10. Between the ages of ___ and ___, did you ever lie to get out of trouble?

(Relevant) 11. Do you have that $20 with you now?

The age used for the probable-lie questions was one year less than the subject's current age to exclude the issue under investigation from consideration.

The test questions for participants assigned to the directed lie condition were the same as those presented to participants in the probable-lie condition, except that the probable-lie questions in positions 4, 7, and 10 were replaced with the following directed lie questions:

(Directed Lie) 4. In your entire life, did you ever tell even one lie?

(Directed Lie) 7. Have you ever broken a rule or regulation?

(Directed Lie) 10. Did you ever make a mistake?

After reviewing the test questions, a probable-lie or directed lie test was administered. The interval between question onsets was at least 25 s, and the interval between repetitions of the question sequence was between one and three minutes. For probable-lie participants, after the first chart, the examiner asked the participant if they noticed anything unusual during the polygraph examination. After the second chart, the examiner asked the participant if they noticed anything unusual when they were asked one of the probable-lie questions. For directed lie participants, after each of the first three charts the examiner asked the participant if they were lying to the directed lie items and if they felt any differently when they lied. These procedures were designed to draw the participant's attention to the comparison questions, and reduce the risk of false positive errors.

The question sequence was presented five times. Neutral and comparison questions were rotated over repeated presentations of the question sequence such that each relevant question was preceded by each neutral and each comparison question at least once. The orders of presentation of the questions were not reviewed with the participant in advance.

At the conclusion of the test, the sensors were removed, and the subject was asked to complete posttest questionnaires. The probability that the participant was truthful was then computed from the physiological responses using algorithms described elsewhere (Kircher & Raskin, 1988). If the probability of truthfulness exceeded 0.70, the participant was paid $80, $30 for their time and a $50 bonus. Otherwise, the participant was paid $30. The participant was then debriefed and released.

Apparatus

The CPS-LAB system (Scientific Assessment Technologies, SLC, UT) was used to configure the data collection hardware, specify storage rates for the physiological signals, and build automated data collection protocols. CPS-LAB also was used to collect, edit, and score the physiological data.

The physiological data acquisition subsystem (PDAS) of CPS-LAB generated analog signals for thoracic and abdominal respiration, skin conductance, cardiograph, finger pulse amplitude, skin potential, and cardiotachometer. In addition, calibrated analog output from a Ohmeda 2300 Blood Pressure Monitor was routed to a general-purpose coupler on the PDAS. Each of the eight analog signals was digitized at 1000 Hz with a Metrabyte DAS 16F analog-to-digital converter installed in a 50 MHz PC compatible 486 computer with 16 MB of RAM.

Respiration was recorded from two Hg strain gauges secured with Velcro straps around the upper chest and the abdomen just below the rib cage. The strain gauge changed in resistance as the subject breathed. Resistance changes were recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 13$Hz.

Skin conductance was obtained by applying a constant voltage of .5V to two UFI 10mm Ag-AgCl electrodes filled with .05M NaCl in a Unibase medium. The electrodes were taped with adhesive collars to the distal phalanx of the ring and last fingers of the left hand. The signal was recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 6$ Hz.

The cardiograph was recorded from a blood pressure cuff wrapped around the right upper arm and inflated to 55 to 60 mm of Hg at the beginning of each chart. The cuff was connected by rubber tubing to a Motorola MPX10DP pressure transducer in the PDAS. The output from the pressure transducer was amplified and recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 8.8$ Hz.

Finger pulse amplitude was obtained from a UFI photoplethysmograph attached to the first finger of the left hand with a Velcro strap. The signal from the photocell was AC-coupled with a 0.2-second time constant and a 2-pole, low-pass filter, $f_c = 10$ Hz.

The electrocardiogram was obtained from Lead II using disposable, pre-gelled Red Dot™ Ag-AgCl snap electrodes taped to the left arm and right leg. The PDAS generated a 20 ms square wave pulse that coincided with the R-wave in the electrocardiogram. The square wave from the PDAS was routed to the analog-to-digital converter, and the CPS-LAB software measured and stored the time between successive pulses (interbeat intervals).

Skin potential was recorded from Beckman 10mm Ag-AgCl electrodes filled with .05 M NaCl in a Unibase medium attached to the thumb of the left hand (active site) and the lower arm, just below the elbow (inactive site). The inactive site was rubbed with alcohol prior to applying the electrode. Skin potential was recorded DC-coupled with a 2-pole, low-pass filter, $f_c = 10$ Hz. A 39.2 K ohm resistor was soldered in series with the reference (inactive) electrode to prevent variations in skin potential from affecting the skin conductance recordings.

The finger cuff of the Finapres Blood Pressure Monitor was attached with Velcro to the middle phalanx of the middle finger on the left hand. Continuous calibrated voltage changes from the Finapres Monitor were routed to a general purpose coupler on the PDAS where it was recorded

DC-coupled with a 2-pole, low-pass filter, $f_c = 10$ Hz. The voltage changes were converted to absolute blood pressure in mm of Hg.

The 1000 Hz samples for each channel were reduced prior to storing them on the hard disk by averaging the samples for successive epochs. Respiration and electrodermal channels were stored at 10 Hz. Cardiograph, finger pulse, and blood pressure signals were stored at 100 Hz. The cardiotachometer produced an interbeat interval measured to the nearest ms for each heart beat.

Calibration Procedures

To assess the relationships between tonic arousal and phasic reactivity to test questions, it was necessary to convert the raw data in analog-to-digital converter units to absolute units for skin conductance, skin resistance, and blood pressure. The CPS-LAB system already provided the interbeat intervals in ms that were required to study heart period and vagal tone. For skin conductance, a separate multiple regression equation was developed for each of six possible gain settings on the PDAS. Each equation predicted known conductances from the offset on the front panel, internal PDAS digital-to-analog (DAC) offset settings, and observed analog-to-digital converter values. The conductance values used to calibrate the instrument ranged from 1 μSiemen (1 M ohm) to 50 μSiemens (20 K ohm). External (front panel) and internal (DAC) offsets were also systematically varied to ensure that the resulting equation would work for any configuration of gain and offset settings. Each equation accounted for over 99.8% of the variance in known inputs.

Since resistance (R) is the reciprocal of conductance (G), skin resistance was obtained by inverting the calibrated skin conductance signal prior to extracting measurements of response amplitude, i.e., $R = 1/G$.

The methods used to derive absolute measures of skin conductance were used to develop equations to measure the output voltages generated by the Finapres. The Finapres generated a voltage that ranged from 0V to 2V that was linearly related to blood pressure that ranged from 0 mm Hg to 200 mm Hg. Again, the error of measurement was negligible; the regression equations accounted for over 99.8% of the variance in the voltages obtained from the Finapres.

Measurements of Autonomic Activity

Tonic Arousal

For each autonomic measure, tonic levels were measured by calculating the mean of 5-second epochs of basal activity prior to the onset of each question within each chart. The frequency of nonspecific skin conductance responses was measured in addition to measuring skin conductance level. Nonspecific responses were measured during the last 15 seconds of the recording interval that followed each neutral question.

Preliminary examination of the distributions of all tonic measures of arousal revealed, as expected, significant positive skew for measures of skin conductance level and number of skin conductance responses. Log transformations of the skin conductance measures normalized the distributions (Venables & Christie, 1980). To avoid undefined values (log of zero), a value of 1 was added to each skin conductance score prior to taking the log of the score.

Phasic Reactivity

Measures of phasic reactivity included peak amplitude of the skin conductance response (SCR), peak amplitude of the skin resistance response (SRR), peak amplitude of increases in systolic blood pressure (SBPR), and peak amplitude of increases in diastolic blood pressure (DBPR). For heart period, the mean heart period for the last 2 beats prior to question onset was subtracted from the longest heart period between 4 and 15 sec after question onset (HPR, Podlesny & Kircher, 1999). A log transformation of SCR also was performed to normalize its distribution.

Indices of Differential Reactivity to Comparison and Relevant Questions

Following our standard protocol (Kircher & Raskin, 1988), an index of differential reactivity to comparison and relevant questions was computed for each subject and each autonomic measure. For example, each subject provided 18 measurements of skin conductance amplitude for the three comparison questions and the three relevant questions on each of the first three charts. The 18 measurements were converted to Z scores. The mean of the nine Z scores for relevant questions was then subtracted from the mean of the nine Z scores for comparison questions.

Indices of differential reactivity can be weighed and combined by means of a discriminant function or regression equation to maximize discrimination between truthful and deceptive subjects (Kircher & Raskin, 1988; Kircher, Woltz, Bell, & Bernhardt, 1998). An index of differential reactivity is analogous to the total numerical score assigned by the polygraph examiner for a particular channel. The index was positive when the mean reaction to comparison questions was greater than the mean reaction to relevant questions, and the index was negative when the reactions to relevant questions were greater. Since innocent subjects were expected to react more strongly to comparison questions and guilty subjects were expected to react more strongly to relevant questions, we expected positive scores for innocent subjects and negative scores for guilty subjects.

For all variables except respiration, a large measured response was indicative of a strong reaction. For respiration excursion, suppressed respiratory activity was indicative of a strong reaction. Thus, innocent subjects were expected to show relatively small measured respiration responses (suppression) to comparison questions, whereas guilty subjects were expected to show relatively small measured respiration responses (suppression) to relevant questions. To maintain consistency of interpretation across physiological measures, the sign of the index of differential reactivity for respiration was reversed.

Results

Preliminary Analyses

Treatment-Related Attrition. Thirty-three individuals assigned to the guilty condition (16%) refused to participate after they had received their tape-recorded instructions, whereas none of the innocent subjects declined to participate. Consequently, subjects who agreed to commit the mock crime may have been sampled from a population that differed in certain respects from the more general population from which innocent subjects were drawn. For example, subjects who remained in the guilty condition on average may have been older or less anxious than subjects in the innocent condition. Preliminary tests were conducted to explore the possibility that guilty and innocent groups differed on measures of marital status, ethnicity, occupation, age, education, hours of sleep, the Marlowe-Crowne scale (Crowne & Marlowe, 1964), Rotter Trust scale (Rotter, 1967), and two anxiety scales (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). The guilty and innocent subjects who completed the experiment did not differ significantly on any of the demographic or personality measures.

Effects of Feedback. The design of the original study (DoDPI97-P-0016) included feedback conditions that were not representative of current field practice. Factorial ANOVAs were conducted to determine if the effects of interest varied as a function of Feedback and other facets of the design. The results of the ANOVAs and effect sizes are summarized in Table 1.

Since Question Type (comparison versus relevant) was not a factor in these analyses, no significant effects were expected, with one possible exception. The effect of Guilt obtained for heart period level (HPL) is consistent with data from a field study reported by Krapohl and Ansley (1999). Guilty subjects had significantly shorter heart periods (higher heart rates) than did innocent subjects. In general, however, no more than 4% of the variance in any tonic or phasic measure was related to Feedback or any interaction between Feedback and Guilt, Test Type, or Sex. In addition, the number of significant effects was no greater than what one would expect to occur by chance.

Since an objective of the present research was to assess the *relationship* between tonic arousal and phasic reactivity, preliminary tests also were conducted for heterogeneity of regression slopes among the various treatment conditions. If the relationship between tonic arousal and reactivity varied as a function of Feedback, then the research plan was to analyze data only from the effective feedback groups. The plan was to focus on the effective feedback groups because the procedures used for those groups were most similar to the procedures used in actual field polygraph examinations.

Separate tests for heterogeneity of regression were conducted for skin conductance, skin resistance, systolic and diastolic blood pressure, and heart period. The results revealed no evidence that the slopes of the regression lines that related phasic reactivity to tonic level varied across the 16 cells in the design matrix (see Figure 1). Nor was there evidence that the relationships between differential reactivity and tonic arousal varied across treatment conditions.

Table 1. Proportions of Variance (η^2) in Tonic and Phasic Measures Explained by Guilt, Test Type, Feedback, and Sex

Factor	SCN[1]	SCL[2]	SCR[3]	SRL[4]	SRR[5]	SBPL[6]	SBPR[7]	DBPL[8]	DBPR[9]	HPL[10]	HPR[11]
G-uilt	.00	.01	.01	.01	.00	.00	.02*	.00	.02*	.03*	.01
T-est	.01	.02*	.01	.01	.05*	.00	.01	.00	.02*	.00	.00
F-eedback	.03*	.01	.00	.01	.00	.00	.02	.00	.02	.00	.00
S-ex	.00	.02*	.00	.02*	.03*	.00	.00	.04*	.01	.01	.00
GxT	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
GxF	.00	.01	.00	.00	.01	.01	.00	.01	.02	.01	.01
GxS	.00	.00	.00	.00	.00	.00	.02*	.00	.02*	.00	.01
TxF	.01	.01	.01	.01	.02	.02	.02	.02	.04*	.00	.01
TxS	.00	.00	.00	.00	.01	.00	.00	.00	.01	.00	.00
FxS	.00	.03*	.02	.01	.01	.02	.03*	.03*	.00	.01	.01
GxTxF	.00	.01	.01	.02	.01	.02	.01	.02	.02	.02	.02
GxTxS	.01	.01	.00	.00	.00	.01	.01*	.00	.00	.00	.00
GxFxS	.01	.01	.01	.01	.00	.00	.01	.00	.01	.00	.01
TxFxS	.02	.01	.01	.01	.00	.00	.01	.00	.00	.04*	.02
GxTxFxS	.01	.00	.00	.00	.00	.00	.02	.01	.02	.01	.00

[1] Log Number of Nonspecific Skin Conductance Responses
[2] Log Skin Conductance Level
[3] Log Skin Conductance Response
[4] Skin Resistance Level
[5] Skin Resistance Response
[6] Systolic Blood Pressure Level
[7] Systolic Blood Pressure Response
[8] Diastolic Blood Pressure Level
[9] Diastolic Blood Pressure Response
[10] Heart Period Level
[11] Heart Period Response
* $p < .05$

In summary, the results shown in Table 1 revealed little or no difference among the groups on mean levels of tonic arousal and reactivity. The tests for heterogeneity of regression also revealed no reliable differences among the groups in the relationships between arousal and reactivity. The results of these preliminary analyses suggested that the relationships of interest were the same or similar across the various facets of the design. For example, the relationship between tonic arousal and phasic reactivity for guilty subjects in the effective feedback condition did not differ significantly from the relationship between tonic arousal and phasic reactivity for innocent subjects who received neutral or ineffective feedback.

To maximize power and precision, subsequent analyses were conducted with the entire sample of 336 cases. However, partially redundant, parallel tests were conducted using only the subjects in the effective feedback conditions. These tests were conducted to verify that the effects observed in the entire sample were consistent with those from the subjects who were given the numbers test and received effective feedback (n = 120).

Table 2. Means (and Standard Deviations) of Physiological Measures

Physiological Measure	*Entire Sample (N=336)*		*Probable-lie Effective Feedback (n=60)*		*Directed Lie Effective Feedback (n=60)*	
Skin Conductance Log Number of Responses	.235	(.187)	.203	(.185)	.187	(.173)
Skin Conductance Log Level (μS)	.611	(.238)	.634	(.230)	.557	(.240)
Skin Conductance Log Amplitude (μS)	.144	(.119)	.118	(.106)	.157	(.148)
Skin Conductance Differential Reactivity	-.034	(.792)	-.201	(.876)	-.115	(.825)
Skin Resistance Tonic Level (K ohms)	291.36	(190.24)	277.26	(163.21)	327.01	(195.63)
Skin Resistance Amplitude (K ohms)[b]	-19.65	(19.10)	-13.51	(11.26)	-26.65	(28.77)
Skin Resistance Differential Reactivity	-.041	(.839)	-.236	(.907)	-.131	(.881)
Systolic Blood Pressure Tonic Level (mm of Hg)[a]	103.69	(15.49)	100.37	(17.53)	106.14	(12.86)
Systolic Blood Pressure Amplitude (mm of Hg)[a]	13.56	(4.98)	12.44	(4.38)	14.18	(4.87)
Systolic Blood Pressure Differential Reactivity	.013	(.686)	-.067	(.657)	.012	(.662)
Diastolic Blood Pressure Tonic Level (mm of Hg)	62.7	(10.91)	61.89	(12.85)	64.03	(9.90)
Diastolic Blood Pressure Amplitude (mm of Hg)[b]	7.49	(2.36)	6.66	(1.70)	7.93	(2.20)
Diastolic Blood Pressure Differential Reactivity	.055	(.679)	-.018	(.693)	-.011	(.684)
Heart Period Tonic Level (msec)	760.22	(120.71)	777.95	(120.92)	758.03	(130.84)
Heart Period Phasic Response (msec)	50.18	(33.44)	53.12	(32.98)	50.23	(34.01)
Heart Period Differential Reactivity	-.257	(.661)	-.071	(.718)	-.241	(.644)

[a]The difference between the means for probable-lie and directed lie tests was significant at $p < .05$.
[b]The difference between the means for probable-lie and directed lie tests was significant at $p < .01$.

Table 2 shows the mean and standard deviation of each physiological measure for the entire sample of 366 subjects and for probable-lie and directed lie subjects who received effective feedback. Subjects in the probable-lie condition showed larger skin resistance responses, whereas subjects in the directed lie condition had higher systolic blood pressure levels and produced stronger systolic and diastolic blood pressure responses.

Correlations Between Tonic Arousal and Phasic Reactivity

One objective of the present research was to determine if tonic levels of autonomic arousal are related to the magnitude of physiological reactions to test questions. Table 3 shows the correlation between each tonic measure of autonomic activity and the corresponding mean amplitude of the phasic response for all cases combined and for the two effective feedback conditions.

Table 3. Correlations Between Measures of Tonic Level and Phasic Reactivity

Tonic Arousal	*Entire Sample (N=336)*	*Probable-lie Effective Feedback (n=60)*	*Directed Lie Effective Feedback (n=60)*
Skin Conductance Log Nonspecific Responses	.661 **	.758 **	.717 **
Skin Conductance Log Level	.580 **	.596 **	.610 **
Skin Resistance Level	-.199 **	-.079	-.227
Blood Pressure Systolic Level	.102	.189	.181
Blood Pressure Diastolic Level	.114 *	-.077	.128
Heart Period Level	.138 *	.133	.052

* $\underline{p} < .05$
** $\underline{p} < .01$

As shown in Table 3, the strongest correlations were between measures of tonic and phasic electrodermal activity. For skin conductance, high tonic levels were associated with relatively strong skin conductance responses. The correlation between the log number of nonspecific responses and the log skin conductance level was .57 for the entire sample. Since skin resistance is the inverse of skin conductance, the negative correlation for skin resistance was expected. High tonic levels of skin resistance were associated with relatively small skin resistance responses.

For the entire sample, a small but significant positive correlation also was observed for heart period. Long interbeat intervals were associated with large increases in heart period. Stated differently, low basal heart rates were associated with the greatest drops in heart rate following question onset. High blood pressure levels tended to be positively related to the magnitude of the phasic response, but the correlations were generally small and may not be reliable.

Correlations Between Tonic Arousal and the Criterion (Guilt)

Another objective of the present study was to determine if measures of tonic arousal could be used to improve the accuracy of probable or directed lie polygraph tests. The first method used to assess the potential usefulness of each tonic arousal measure was to calculate a point-biserial correlation between the measure and the criterion. The criterion was a dichotomous variable that distinguished between guilty (coded 0) and innocent subjects (coded 1; Kircher & Raskin, 1988). Table 4 shows these point-biserial correlations for the entire sample and the effective feedback conditions separately.

Table 4. Point-biserial Correlations Between Measures of Tonic Level and the Guilt/Innocence Criterion

Physiological Measure	*Entire Sample (N=336)*	*Probable-lie Effective Feedback (n=60)*	*Directed Lie Effective Feedback (n=60)*
Skin Conductance Log Number of Responses	-.042	.047	.011
Skin Conductance Log Level	-.093	.013	-.089
Skin Resistance Level	.110 *	.057	.136
Systolic Blood Pressure Level	-.106	-.271 *	.045
Diastolic Blood Pressure Level[a]	-.031	-.142	.251
Heart Period Level	.206 **	.424 **	.092

* $p < .05$
** $p < .01$
[a]The difference between the correlations for probable-lie and directed lie effective feedback conditions was significant at $p < .05$.

Most of the correlations with the criterion were not significant. The greatest correlations were obtained from subjects who were given probable-lie tests. In that condition, guilty subjects had higher heart rates (M = 84.3 BPM) than did innocent subjects (M = 73.8 BPM; r = .42). Guilty probable-lie subjects also had higher systolic blood pressure levels (M = 105.1) than did the innocent subjects (M = 95.7; r = -.27). There were no significant differences in tonic arousal between guilty and innocent subjects who received directed lie tests.

Regression Equations that Combine Tonic Level and Standard Measures to Predict the Criterion

Several types of multiple regression analyses were conducted to determine if measures of tonic arousal could be used in combination with traditional indices of differential reactivity to improve discrimination between truthful and deceptive individuals. For each type of analysis, a tonic measure of arousal was added to a base model, and the contribution of the tonic measure to the prediction equation was tested for significance. In all of these analyses, the criterion to be predicted was a dichotomous variable that distinguished between guilty (coded 0) and innocent subjects (coded 1).

The predictor variables in the base model were indices of differential reactivity. The base model contained one index of differential reactivity for each response system. Each base model contained a respiration index, an electrodermal index, and a cardiovascular index. The particular set of base model predictors depended on the tonic measure of arousal that was to be tested for significance. The base model contained an index of differential reactivity from the same channel as the one that provided the measure of tonic arousal. For example, when the tonic measure of arousal was systolic blood pressure level, the index of differential reactivity for systolic blood pressure responses was substituted for the cardiograph in the base model. We used this approach because we expected that tonic systolic blood pressure would correlate more highly with its own corresponding index of differential reactivity than with some other measure of cardiovascular reactivity.

The stronger the correlation between measures of tonic arousal and differential reactivity, the more likely it was that the tonic measure would serve as a suppressor variable and would make a significant contribution to the prediction equation. Suppression occurs when one variable in the model (tonic level) is uncorrelated with the criterion, but it is highly correlated with another variable (differential reactivity index) that *is* correlated with the criterion (Cohen & Cohen, 1975). A suppressor variable improves the diagnostic validity of a predictor variable by removing noise from the predictor variable that attenuates the predictor variable's correlation with the criterion. In general, the best suppressor variable is one that correlates near zero with the criterion, yet it is highly correlated with another variable in the regression equation.

Linear Analysis of Tonic Measures of Arousal. Table 5 summarizes the results from the first set of regression analyses. The first column of Table 5 lists the indices of differential reactivity included in the base model. The second column shows the tonic measure of arousal that was added to the base model. The third column shows the proportion of variance in the criterion explained by the base model for the entire sample of 336 subjects (R^2). The fourth column shows the increment in the proportion of variance explained by the measure of tonic arousal (ΔR^2). The last four columns show the R^2 and the ΔR^2 for probable-lie and directed lie subjects who received effective feedback.

The results in Table 5 indicate that the various measures of tonic arousal added little to the regression equations. The only measure of tonic arousal to make a significant contribution to the base model was heart period level. Discriminant analyses were performed to compare the accuracy of dichotomous classifications into truthful and deceptive groups for the base model and the model that included heart period level. Although the contribution of heart period level to the regression equation was statistically significant, the classification accuracy was slightly lower for the model that included heart period level (80.1%) than for the base model (81.5%).

Table 5. Proportions of Variance in the Criterion Explained by Indices of Differential Reactivity (R^2) and Increments in Variance Explained by Tonic Measures of Arousal (ΔR^2)

		Entire Sample (N = 336)		*Probable-lie Effective Feedback (n = 60)*		*Directed Lie Effective Feedback (n = 60)*	
Base Model	**Tonic Measure**	R^2	ΔR^2	R^2	ΔR^2	R^2	ΔR^2
SC Amplitude CP Amplitude Respiration Length	SC Level	.406**	.002	.633**	.019	.404**	.010
SC Amplitude CP Amplitude Respiration Length	SC Number of Responses	.406**	.001	.633**	.020	.404**	.006
SR Amplitude CP Amplitude Respiration Length	SR Level	.406**	.003	.622**	.004	.421**	.010
SC Amplitude SBP Amplitude Respiration Length	SBP Level	.423**	.002	.618**	.009	.423**	.004
SC Amplitude DBP Amplitude Respiration Length	DBP Level	.420**	.000	.597**	-.005	.415**	.028
SC Amplitude CP Amplitude Respiration Length HP Increase	HP Level	.408**	.019**	.633**	.014	.408**	.024

** $p < .01$

Nonlinear Analysis of Tonic Measures of Arousal. Indices of differential reactivity reflect the extent to which subjects responded more strongly to comparison or relevant questions. Although tonic measures of electrodermal and cardiovascular activity were linearly related to the magnitude of phasic reactions to test questions (see Table 3), tonic measures may be nonlinearly related to the *differences* in reactions to comparison and relevant questions. Subjects who react strongly to comparison questions and have large positive indices of differential reactivity may have relatively high tonic levels of activity. Similarly, subjects who react strongly to relevant questions and have large negative indices of differential reactivity also may have high tonic levels of activity. Conversely, subjects who show little difference in their reactions to comparison and relevant questions may have relatively low levels of tonic activity. Under these conditions, a nonlinear transformation of tonic arousal may provide a measure that removes noise from the corresponding index of differential reactivity more effectively than the untransformed measure of tonic arousal.

To explore this possibility, the observed sign of the difference between comparison and relevant questions was applied to the tonic measure prior to adding the tonic measure to the regression equation. Given a strong nonlinear relationship between tonic arousal and differential reactivity, large positive indices of differential reactivity would be associated with large positive measures of tonic arousal, and large negative indices of differential reactivity would be associated with large negative measures of tonic arousal. That is, the transformation would result in a strong linear relationship between tonic arousal and differential reactivity. As noted above, for suppression to occur, there should be a strong correlation between the suppressor variable and the predictor variable.

Table 6 shows the results of the multiple regression analyses when the transformed values of tonic arousal were included in the regression equation to predict the criterion. Since none of the increments in the proportion of variance explained was significant, there was no evidence that any transformed measure of tonic arousal increased the diagnostic validity of the base model.

Table 6. Proportions of Variance in the Criterion Explained by Indices of Differential Reactivity (R^2) and Increments in Variance Explained by Transformed Tonic Measures of Arousal (ΔR^2)

		Entire Sample (N = 336)		*Probable-lie Effective Feedback (n = 60)*		*Directed Lie Effective Feedback (n = 60)*	
Base Model	**Tonic Measure**	**R^2**	**ΔR^2**	**R^2**	**ΔR^2**	**R^2**	**ΔR^2**
SC Amplitude CP Amplitude Respiration Length	SC Level	.406**	.001	.633**	.003	.404**	.000
SC Amplitude CP Amplitude Respiration Length	SC Number of Responses	.406**	.003	.633**	.000	.404**	.001
SR Amplitude CP Amplitude Respiration Length	SR Level	.406**	.007	.622**	.000	.421**	.009
SC Amplitude SBP Amplitude Respiration Length	SBP Level	.423**	.005	.618**	.002	.423**	.000
SC Amplitude DBP Amplitude Respiration Length	DBP Level	.420**	.002	.617**	.004	.415**	.033
SC Amplitude CP Amplitude Respiration Length HP Increase	HP Level	.408**	.007	.633**	.008	.408**	.017

** $p < .01$

Analysis of Covariance. A final attempt was made to test if tonic measures of skin conductance can be used in combination with indices of differential reactivity to improve the accuracy of polygraph outcomes. When the raw measurements of skin conductance response amplitude for a given individual are transformed to z scores, the resulting z scores have a mean of zero and unit variance. Since the z scores for all subjects have the same mean and the same variance, there are no differences among subjects in the average magnitude of skin conductance responses to test questions. Another method for removing differences among subjects is to use the original raw measurements of skin conductance amplitude and include the tonic level as a covariate in the regression equation.

Two measures of tonic skin conductance activity were derived. Therefore, two regression equations were created for each sample. The equations included raw differences between comparison and relevant response amplitude (in μS) and either log skin conductance level or log number of spontaneous skin conductance responses as the covariate. Table 7 summarizes the results of the regression analyses. To permit comparisons with the traditional approach, the regression equations that included indices of differential reactivity for skin conductance, cardiograph, and respiration based on differences between z scores are shown in the top row of Table 7. These results also appear along the top row of Table 6.

In the entire sample, all variables contributed significantly to their respective regression equations. As expected, the regression coefficients for both SC covariates (SC level and SC number of responses) were significant. These findings suggest that the covariates functioned as predicted; they removed variance among individuals in the raw differences in SC responses to comparison and relevant questions. However, the overall proportions of variance in the criterion explained by the current model with z-transformed measures of SC amplitude (R^2 = .406, .633, and .404) were generally greater than those obtained for either of the proposed models with raw measures of SC amplitude. Therefore, there was no advantage in using tonic SC level or the number of spontaneous responses as a covariate in the regression equation.

Although the proportions of variance explained by the regression models for the probable-lie test were generally greater than those for the directed lie test, none of the differences were statistically significant. Differences between probable-lie and directed lie tests are discussed in the final report for the original study (DoDPI97-P-0016).

Table 7. Regression Results for Current Model with Standardized Measurements and Alternative Models with Absolute Measures of SCR Amplitude and Tonic Activity

	Entire Sample (N = 336)		*Probable-lie Effective Feedback (n = 60)*		*Directed Lie Effective Feedback (n = 60)*	
Model	**B**	**R^2**	**B**	**R^2**	**B**	**R^2**
Standardized Measures with No Covariates		.406**		.633**		.404**
Standardized SCR Amplitude	.550**		.608**		.591**	
Standardized CPR Amplitude	.099*		.141		.086	
Standardized Respiration Excursion	.176**		.293**		-.025	
SC Level as a Covariate		.341**		.515**		.380**
SCR Amplitude (in μS)	.186**		.174		.128	
Standardized CPR Amplitude	.195**		.241*		.140	
Standardized Respiration Excursion	.187**		.301**		-.007	
SC Level (in μS)	.295**		.325*		.440*	
Number of Spontaneous SC Responses as a Covariate		.325**		.474**		.335**
SCR Amplitude (in μS)	.231**		.274*		.293	
Standardized CPR Amplitude	.221**		.308**		.209	
Standardized Respiration Excursion	.183**		.286**		-.028	
SC Number of Responses	.226**		.154		.220	

*Differed from 0, p < .05.
**Differed from 0, p < .01.

Discussion

One objective of the present research was to assess relationships between tonic levels of arousal and the magnitude of phasic reactions to test questions. Measures of tonic arousal were obtained from 336 subjects who participated in a previous experiment (DODPI97-P-0016). That experiment was designed to assess the effects of the demonstration test on the accuracy of subsequent probable-lie and directed-lie polygraph examination. It also assessed the effects of various types of feedback concerning the outcome of the demonstration test on the accuracy of the subsequent polygraph examinations. Preliminary tests for heterogeneity of regression revealed no evidence that the strength or direction of relationships between tonic arousal and phasic reactivity varied over treatment conditions. Analyses of group means revealed several significant effects of Feedback and interactions of Feedback with other factors on selected physiological measures. However, the effects were generally small, and they did not affect the nature of the relationships between tonic and phasic measures of autonomic activity.

These preliminary tests justified an analysis of all subjects. Nevertheless, in addition to analyzing the entire sample, we conducted parallel analyses of the probable-lie and directed lie groups that had been given numbers tests and effective feedback. The two effective feedback groups were analyzed separately because the procedures for those groups were most similar to the procedures used by field polygraph examiners.

Six measures of tonic arousal were derived. They included the number of nonspecific skin conductance responses, skin conductance level, skin resistance level, systolic and diastolic blood pressure levels, and heart period level. The results obtained for the entire sample were similar to those obtained for the effective feedback subgroups. For the entire sample, five of the six measures of tonic arousal correlated significantly with the mean magnitude of phasic responses to test questions. However, the correlations for the cardiovascular measures were generally small and were statistically significant only for the entire sample. This suggests that there may be small but reliable relations between tonic and phasic measures of cardiovascular activity, but the power was too low to provide evidence of these relationships in the effective feedback subgroups.

Phasic skin conductance responses correlated highly with the number of nonspecific skin conductance responses ($r > .66$) and the skin conductance level ($r > .57$). These findings are consistent with those reported by Hord et al. (1964) and by Schell, Dawson, and Filion (1988). Hord et al. reported correlations between skin conductance level and skin conductance responses that ranged from .35 to .77.

Skin resistance was derived from measured skin conductance activity. The expected negative correlation between tonic skin resistance and skin resistance was significant for the entire sample. However, the correlation was considerably smaller in magnitude ($r = -.20$) than the correlations obtained for skin conductance, and the correlation was not significant for either of the effective feedback groups. These results suggest that skin resistance responses are less dependent on basal levels of activity than are skin conductance responses.

Hord et al. (1964) also reported strong negative correlations between heart rate level and increases in heart rate. In their study, subjects with high tonic heart rates showed the smallest increases in heart rate. In the present study, heart period level was positively correlated with increases in heart period but only marginally so. Since heart period and heart rate are inversely related, a positive correlation between measures of tonic and phasic heart period was expected. However, the magnitude of the correlation obtained in the present study ($r = .14$) was considerably less than the correlations reported by Hord et al. ($r = -.36$ to $-.64$). This discrepancy may be due in part to the nonlinear nature of the relationship between heart period and heart rate. To derive heart rate (in BPM) from heart period (in seconds), one finds the reciprocal of the heart period and multiplies by 60. Although there is a strong negative relationship between heart rate and heart period, it is not perfect. When we reexamined the tonic-phasic relationship after transforming our measures of heart period to heart rate, the strength and sign of the correlation changed from $r = .14$ to $r = -.21$. The sign of the correlation for heart rate was consistent with Hord et al., and the magnitude of the correlation increased. However, the observed relationship was still substantially weaker than those reported by Hord et al..

The present study also differed from the study by Hord et al. in how the cardiac response was defined. Hord et al. asked subjects to listen for the onset of a tone and measured *increases in heart rate*. In contrast, we measured *increases in heart period*, which correspond to decreases in heart rate. Measures of cardiac deceleration are traditionally used in research on polygraph techniques (Podlesny & Kircher, 1999; Raskin, 1979) and are more diagnostic than measures of cardiac acceleration (Kircher & Raskin, 1988). The drop in heart rate in response to test questions with signal value may be indicative of an orienting response (Graham & Clifton, 1966), or it may be a reflexive response to a rapid rise in blood pressure (Raskin, 1979).

The availability of continuous measures of blood pressure and the large sample size provided an opportunity to detect even small correlations between tonic levels of blood pressure and phasic blood pressure reactions to test questions. The correlation between tonic diastolic blood pressure and the phasic diastolic blood pressure response was statistically significant for the entire sample. However, the correlation was small ($r = .11$), and it was not significant for either effective feedback group. In addition, the sign of the correlation was negative for one subgroup and was positive for the other. There was no evidence of a relationship between tonic and phasic systolic blood pressure responses. These results suggest that if there is a linear relationship between tonic and phasic blood pressure activity during polygraph examinations, then it is a small one and, for all practical purposes, may be ignored.

Current computer and numerical scoring methods for analyzing polygraph charts do not consider individual differences in tonic levels of physiological activity. Decisions are based exclusively on differences between phasic reactions to comparison and relevant questions. Not only are individual differences in tonic arousal ignored, at least two computer programs systematically remove differences among individuals in phasic responses as well (Kircher & Raskin, 1988; Olsen et al., 1997). These algorithms transform raw measurements of reactions to comparison and relevant questions to standard scores. The transformation to standard scores is a linear one so that a subject who reacts more strongly to a particular type of question (e.g., relevant) will continue to show proportionally stronger reactions to that type of question. However, since the standard scores for all subjects have the same mean ($M_Z = 0$) and variance ($S^2 = 1$), the transformation removes all differences among subjects in the mean magnitude and variance of responses to the test questions. Although the advantages of within-subject standardization and within-subject comparisons of reactions to comparison and relevant questions have been documented (e.g., Kircher & Raskin, 1981; 1988; Podlesny & Kircher, 1999; Raskin, Kircher, Honts, & Horowitz, 1988), it was not known if additional diagnostic information could be extracted from absolute levels of tonic arousal. The primary purpose of the present study was to explore that possibility.

Three of the six measures of tonic autonomic arousal distinguished between guilty and innocent subjects. Consistent with expectations, in the entire sample, skin resistance was lower and heart rates were higher for guilty subjects than for innocent subjects. Together, these results suggest that guilty subjects were more highly aroused during their polygraph examinations than were innocent subjects. However, the effect on skin resistance was small ($r_{pb} = .11$) and was not significant for the probable-lie or directed lie subsamples. In addition, the effect on heart period was substantial only for probable-lie subjects ($r_{pb} = .42$). For directed lie subjects, the effect was not significant ($r_{pb} = .09$). A third, marginally significant difference between guilty and innocent subjects in the probable-lie condition was obtained for systolic blood pressure level ($r = -.27$). The guilty subjects in that condition had higher tonic systolic blood pressure than did innocent subjects.

The observed, weak relationships between measures of tonic arousal and the criterion were not surprising. There are large differences among individuals in levels of basal autonomic activity (Sternbach, 1966). Having committed a mock crime, subjects' tonic levels of autonomic activity increased. However, with the exception of heart rate for probable-lie subjects, the increases were small relative to the wide range of individual differences inherent in these measures.

Despite the small effects of deception on tonic measures of arousal, it was still possible that they would improve the accuracy of polygraph decisions when combined with traditional measures of differential reactivity. Multiple regression was used to test for an increase in the proportion of criterion variance explained when measures of tonic activity were used in combination with measures of differential reactivity. The proportion of criterion variance explained and the expected proportion of correct decisions are monotonically related. However, the proportion of criterion variance explained was chosen as the primary outcome measure because it is a more sensitive index of predictive validity than the accuracy of decisions.

Initially, a regression equation was created using only measures of differential reactivity to comparison and relevant questions. A measure of tonic arousal was then added to the equation, and its contribution to the proportion of criterion variance explained was tested for significance. To facilitate interpretation of the results, only one tonic measure was added to the regression equation at a time.

A tonic measure of arousal could contribute to the regression equation in one of two ways: as a *predictor* variable or as a *suppressor* variable (Cohen & Cohen, 1975). In general, a useful *predictor* variable accounts for variance in the criterion that is not already explained by other variables in the regression equation. The most useful predictor variables are those that are highly correlated with the criterion and are relatively independent of other variables in the equation.

A *suppressor* variable, on the other hand, contributes to the regression equation by removing (suppressing) a portion of variance from a predictor variable that is unrelated to the criterion. A suppressor variable filters noise from predictor variable and contributes to the regression equation indirectly, by increasing the correlation between the predictor variable and the criterion. The most useful suppressor variables are those that are uncorrelated with the criterion but are highly correlated with one or more predictor variables.

Except for heart period, none of the tonic measures correlated with the criterion. Therefore, there was little hope that any of the tonic measures would serve to predict the criterion directly. There was, however, still some chance that one or more tonic measures of arousal would contribute to the regression equation indirectly by serving as suppressor variables.

Three attempts were made to determine if any of the six tonic measures of arousal would contribute significantly to a base model. The base model contained indices of differential reactivity to comparison and relevant questions for electrodermal, cardiovascular, and respiration channels. The composition of the base model was varied to ensure that each tonic measure was added to a regression model that contained an index of differential reactivity derived from the same channel. For example, when testing the contribution of tonic skin conductance level, the electrodermal index of differential reactivity in the base model was derived from skin conductance responses. Conversely, when testing the contribution of tonic skin resistance level, the electrodermal index of differential reactivity was based on skin resistance responses. This approach was based on the assumption that a tonic measure of arousal would correlate more highly with an index of differential reactivity derived from the same physiological signal than from any other signal. By maximizing the correlation between tonic arousal and an index of differential reactivity, we hoped to increase the chances that the tonic measure would suppress variance in the index of differential reactivity and improve its ability to predict the criterion.

In the first set of analyses, the original measures of tonic activity were entered into the regression equation. In the second set of analyses, transformed measures of tonic activity were entered into the regression equation. Tonic measures were transformed by assigning the sign of the associated index of differential reactivity to the tonic measure before adding the tonic measure to the regression equation. The final set of analyses tested only the contributions of tonic skin conductance level and the number of spontaneous skin conductance responses. In those analyses, differences between the reactions to comparison and relevant questions measured in the original units (μS) were used in the regression equations rather than differences between standardized measurements of reactions to comparison and relevant questions.

The results of these tests were generally disappointing. In only one case did the measure of tonic arousal contribute significantly to the proportion of variance explained. In the analysis of the entire sample of 336 subjects, heart period contributed significantly to the prediction equation. However, heart period accounted for less than 2% of the variance in the criterion and did not improve the accuracy of decisions.

Conclusions

It is difficult to draw firm conclusions from null results when the sample size is small. However, when the sample size is as large as it was in the present study, estimates of population parameters are quite stable. Under these circumstances, it may be concluded that if there are any benefits in using tonic measures of arousal for diagnosing truth and deception, they are likely to be minimal. Based on the present results, there appears to be little need to call on polygraph manufacturers to develop instrumentation that is capable of providing absolute measures of tonic physiological activity.

It is important to note that the data for the present study were obtained from subjects who participated in a mock crime experiment. Whether these laboratory results are representative of the field is an open question. Truthful and deceptive field suspects may show higher levels of tonic arousal than subjects in a laboratory experiment. Indeed, differences in tonic arousal may mediate the observed variance among laboratory studies in the accuracy of polygraph outcomes (Kircher, Horowitz, & Raskin, 1988). However, if the effects of setting were constant for truthful and deceptive subjects, one would not expect measures of tonic arousal to be any more useful in the field than in the lab. We previously compared data from our lab and field polygraph studies and found that differences between comparison and relevant questions in the field sample were generally shifted in the negative direction (Kircher, Raskin, Honts, & Horowitz, 1994). The truthful and deceptive field suspects appeared more deceptive on their polygraph tests than did the truthful and deceptive laboratory subjects. Although the differences between comparison and relevant questions were more negative, the separation between truthful and deceptive individuals was similar for the lab and field samples. The various indices of differential reactivity were as diagnostic for the field sample as they were for the lab sample. In addition, the variances and covariances among various indices of differential reactivity in the lab and field samples were indistinguishable. Thus, the field suspects may have been more physiologically aroused than the laboratory subjects, but this increase arousal had no discernable effect on the diagnosticity of the physiological measures or their interrelationships. The findings from the Kircher et al. study suggest that our laboratory procedures produce relationships between indices of differential reactivity and the criterion and among indices of differential reactivity that closely approximate the relationships that occur in the field. To the extent that they do, the present results suggest that tonic measures of arousal would not prove useful for field applications.

References

Bell, B.G., Raskin, D.C., Honts, C.R., & Kircher, J.C. (1999). The Utah numerical scoring system. *Polygraph, 28,* 1-9.

Boucsein, W. (1992). *Electrodermal Activity.* New York: Plenum Press.

Cohen, J. & Cohen, P. (1975). *Applied multiple regression/ correlation analysis for the behavioral sciences.* New York: Wiley.

Crowne, D. P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence.* New York: Wiley.

Graham, F. K. & Clifton, R. K. (1966). Heart rate change as a component of the orienting response. *Psychological Bulletin, 65,* 305-320.

Hord, D. J., Johnson, L.C. & Lubin, A. (1964). Differential effect of the law of initial values (LIV) on autonomic variables. *Psychophysiology, 1,* 79-87.

Katkin, E.S. (1975). Electrodermal lability: A psychophysiological analysis of individual differences in response to stress. In I. G. Sarason & C. D. Spielberger (Eds.) *Stress and anxiety* (Vol.2). Washington, D. C.: Hemisphere.

Kircher, J. C., Raskin, D. C., Honts, C. R., & Horowitz, S. W. (1994). Genereralizability of statistical classifiers for the detection of deception. *Psychophysiology, 31,* S73. (Abstract)

Kircher, J. C., & Raskin, D. C. (1981). Computerized decision-making in the detection of deception. *Psychophysiology, 18,* 204-205. (Abstract)

Kircher, J.C. & Raskin, D.C. (1988). Human versus computerized evaluations of polygraph data in a laboratory setting. *Journal of Applied Psychology, 73,* 291-302.

Kircher, J.C., Woltz, D.J., Bell, B.G. & Bernhardt, P.C. (1998). Effects of audiovisual presentations of test questions during relevant-irrelevant polygraph examinations and new measures. Final report to the Central Intelligence Agency (Grant No. 110200-997-MO). Salt Lake City: University of Utah, Department of Educational Psychology.

Krapohl, D. & Ansley, N. (1999, August) *Prevalence of diagnostic features in polygraph tracings.* Paper presented at the annual meeting of the American Polygraph Association, Dallas, TX.

Olsen, D. E., Harris, J. C., Capps, M. H., & Ansley, B. A. (1997). Computerized polygraph scoring system. *Journal of Forensic Sciences, 41(1),* 61-70.

Podlesny, J. A. & Kircher, J. C. (1999). The Finapres (volume clamp) recording method in psychophysiological detection of deception examinations: Experimental comparison with the cardiograph method. Final report to the National Institute of Justice (91-IJ-R-026). Laboratory Division, Federal Bureau of Investigation.

Raskin, D. C. (1979). Orienting and defensive reflexes in the detection of deception. In K.D. Kimmel, E. H. Van Olst & J. F. Orlebeke (Eds.) *The orienting reflex in humans.* Hillsdale: Lawrence Erlbaum Associates.

Raskin, D. C. (1989). Polygraph techniques for the detection of deception. In D. C. Raskin (Ed.) *Psychological methods in criminal investigations and evidence* (pp 247-296). New York: Springer Publishing.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35,* 651-665.

Rushmer, R. F. (1976). *Organ physiology: Structure and function of the cardiovascular system (2nd Ed).* Philadelphia: Saunders.

Schell, A. M., Dawson, M. E., & Filion, D. L. (1988). Psychophysiological correlates of electrodermal lability. *Psychophysiology, 25,* 619-632.

Spielberger, C. D., Gorsuch, R.L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA.: Consulting Psychologists Press.

Sternbach, R. A. (1966). *Principles of psychophysiology.* New York: Academic Press.

Venables, P. H. & Christie, M. J. (1980). Electrodermal activity. In I. Martin & P. H. Venables (Eds.) *Techniques in psychophysiology.* New York: John Wiley & Sons.

Wilder, J. (1967). Stimulus and response: The law of initial value. Bristol: Wright.